The Gradient and the Directional Derivative.

If $z = f(x, y)$ is a function of two variables, $\dfrac{\partial f}{\partial x}$ is the rate of change of f in the x-direction and $\dfrac{\partial f}{\partial y}$ is the rate of change of f in the y-direction.

What is the rate of change of f in an *arbitrary* direction? This rate of change should depend on *where we are* and *in what direction* we're moving.

Think of standing on the side of a hill. The rate of change — the slope — certainly depends on where on the hill you're standing. But even if you know where you're standing, the rate of change depends on the direction you move in. If you move uphill, the rate of change is large and positive; If you move downhill, the rate of change is large and negative. If you move along the hill "sideways", the rate of change is 0, because your altitude doesn't change.

We can say "where we are" by giving a *point*; we can say "what direction we're moving in" by giving a *vector*.

Suppose, then, that we have a function f with multiple inputs but only one output. How rapidly is f changing at a point p in the direction of the vector $\vec{v}$?

We use the same procedure that we use to define the ordinary derivative: Move a little bit, measure the average change, then take the limit as the amount you move goes to 0. Here, then, is the definition of the **directional derivative** of f at p in the direction of $\vec{v}$:

$$Df_{\vec{v}}(p) = \lim_{h \to 0} \frac{f(p + h\vec{v}) - f(p)}{h|\vec{v}|}.$$

Notice that $h|\vec{v}|$ is just the amount we moved to get from p to $p + h\vec{v}$.

Let's clean this up, and also obtain a formula which is better for computation. Let $s = h|\vec{v}|$, so $h = \dfrac{s}{|\vec{v}|}$. As $h \to 0$, $s \to 0$, so

$$Df_{\vec{v}}(p) = \lim_{s \to 0} \frac{f\left(p + s\frac{\vec{v}}{|\vec{v}|}\right) - f(p)}{s}.$$

This formula is a little better. Notice that $\dfrac{\vec{v}}{|\vec{v}|}$ is a *unit* vector. This makes sense, because the only function of $\vec{v}$ is to "point out" a direction. We wouldn't want a bigger $\vec{v}$ to give a bigger rate of change. (If you ask someone "Which way is the museum?" and he points toward the museum, does the distance to the museum depend on the length of the finger he points with?)

However, we can use the Chain Rule to do better. Let

$$\vec{\sigma}(s) = p + s\frac{\vec{v}}{|\vec{v}|}.$$

Notice that $\sigma(0) = p$. We have

$$Df_{\vec{v}}(p) = \lim_{s \to 0} \frac{f(\vec{\sigma}(s)) - f(\vec{\sigma}(0))}{s} = (f \cdot \vec{\sigma})'(0) = D(f \cdot \vec{\sigma})(0) = Df(\vec{\sigma}(0)) \cdot \vec{\sigma}'(0).$$

The derivative Df of a function with one output is usually referred to as the **gradient** of f. It is denoted by ∇f. Our formula becomes

$$Df_{\vec{v}}(p) = \nabla f(p) \cdot \vec{\sigma}'(0).$$

However, $\vec{\sigma}'(s) = \dfrac{\vec{v}}{|\vec{v}|}$, so $\vec{\sigma}'(0) = \dfrac{\vec{v}}{|\vec{v}|}$. Here's our formula:

$$Df_{\vec{v}}(p) = \nabla f(p) \cdot \frac{\vec{v}}{|\vec{v}|}.$$

To use the formula, we compute the gradient and the length of $\vec{v}$ and plug in. Before doing some examples, let me say a little more about the gradient.

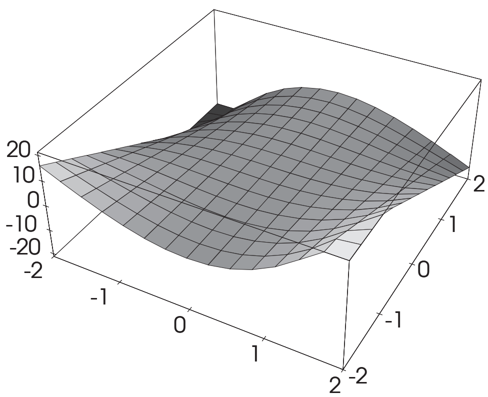
If $w = f(x_1, x_2, x_3, \ldots, x_n)$ is a function with n inputs and one output, Df is a $1 \times n$ matrix — an n-dmensional vector. It's made up of the partials of f (or w) with respect to x_1, x_2, x_3, …, x_n:

$$\nabla f = \left\langle \frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \frac{\partial f}{\partial x_3}, \ldots, \frac{\partial f}{\partial x_n} \right\rangle.$$

For example, if $z = f(x, y)$ is a function of two variables, the gradient is

$$\nabla f = \left\langle \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right\rangle.$$

Example. The picture below shows that graph of $z = x^3 - 3x^2y$. This surface is called the **monkey saddle**. (Why?)



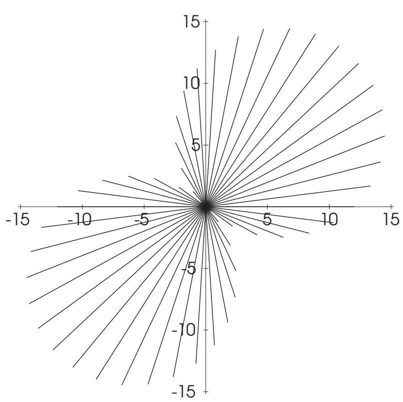
The gradient of z at $(2, 2)$ is $\nabla z(2, 2) = \langle -12, -12 \rangle$. The directional derivative at $(2, 2)$ in the direction of a vector $\vec{v}$ is the rate of change at $(2, 2)$ in the direction of $\vec{v}$.

Since

$$Dz_{\vec{v}}(2, 2) = \nabla z(2, 2) \cdot \frac{\vec{v}}{|\vec{v}|},$$

we can think of the directional derivative as being obtained from the gradient by "scaling down" the gradient in the direction of $\vec{v}$. Here "scaling down" means that we take the scalar projection of $\nabla z(2, 2)$ on $\vec{v}$.

The picture below illustrates this idea. Each segment in the picture below has been scaled so its length is the magnitude of the directional derivative in the direction in which the segment points.



Notice that the largest segments are in the directions $\langle -1, -1 \rangle$ and $\langle 1, 1 \rangle$ — the directions of the gradient and its negative, respectively. The segment shrink in the perpendicular directions, because the rates of change perpendicular to the gradient — i.e. along level curves — are 0. $\square$

Example. Find the rate of change of $f(x,y,z) = xy - yz + xz$ at the point $(1,-2,-2)$ in the direction toward the origin. Is f increasing or decreasing in this direction?

First, compute the gradient at the point:

$$\nabla f = \langle y+z, x-z, -y+x \rangle, \quad \nabla f(1,-2,-2) = \langle -4, 3, 3 \rangle.$$

Next, determine the direction vector. We're at $P(1,-2,-2)$ and we're looking toward the origin $Q(0,0,0)$. We have

$$\overrightarrow{PQ} = \langle -1, 2, 2 \rangle.$$

Make this into a unit vector by dividing by its length:

$$\frac{\overrightarrow{PQ}}{|\overrightarrow{PQ}|} = \frac{1}{3}\langle -1, 2, 2 \rangle.$$

Finally, take the dot product of the unit vector with the gradient:

$$Df_{\vec{v}}(1,-2,-2) = \nabla f(1,-2,-2) \cdot \frac{\overrightarrow{PQ}}{|\overrightarrow{PQ}|}\langle -4, 3, 3 \rangle \cdot \frac{1}{3}\langle -1, 2, 2 \rangle = \frac{16}{3}.$$

f is increasing in this direction, since the directional derivative is positive. $\square$

Example. Calvin Butterball sits in his go-cart on the surface

$$z = x^3 - 2x^2y + x^2 + xy^2 - 2y^3 + y^2$$

at the point $(1,1,0)$. If his go-cart is pointed in the direction of the vector $\vec{v} = \langle 15, -8 \rangle$, at what rate will it roll downhill?

The rate at which he rolls is given by the directional derivative. We have

$$\nabla f = \langle 3x^2 - 6xy + 2x + y^2, -2x^2 + 2xy - 6y^2 + 2y \rangle, \text{ and } \nabla f(1,1) = \langle 0, -4 \rangle.$$

Since $|\langle 15, -8 \rangle| = 17$, we have

$$Df_{\vec{v}}(1,1) = \langle 0, -4 \rangle \cdot \frac{\langle 15, -8 \rangle}{17} = \frac{32}{17} \approx 1.88235. \square$$

Next, we'll take a closer look at the gradient ∇f. Consider the directional derivative formula

$$Df_{\vec{v}}(p) = \nabla f(p) \cdot \frac{\vec{v}}{|\vec{v}|}.$$

The dot product on the right can be written as

$$\nabla f(p) \cdot \frac{\vec{v}}{|\vec{v}|} = |\nabla f(p)| \left| \frac{\vec{v}}{|\vec{v}|} \right| \cos\theta,$$

where θ is the angle between $\nabla f(p)$ and $\vec{v}$.

$\cos\theta$ — and hence, the dot product — will be largest when $\theta = 0$ (so $\cos\theta = 1$). That means that $\vec{v}$ points in the same direction as the gradient — or, to put it another way, the gradient points in the direction in which the rate of change is largest.

1. The gradient points in the direction of most rapid increase of f. For example, the graph of a function of two variables $z = f(x, y)$ is a surface in 3 dimensions. In this case, ∇f is like a compass whose needle always points in the "uphill direction".

   We saw that when $\vec{v}$ points in the direction of the gradient,

$$Df_{\vec{v}}(p) = |\nabla f(p)| \left|\frac{\vec{v}}{|\vec{v}|}\right| \cos\theta = |\nabla f(p)| \cdot 1 \cdot 1 = |\nabla f(p)|.$$

   That is:

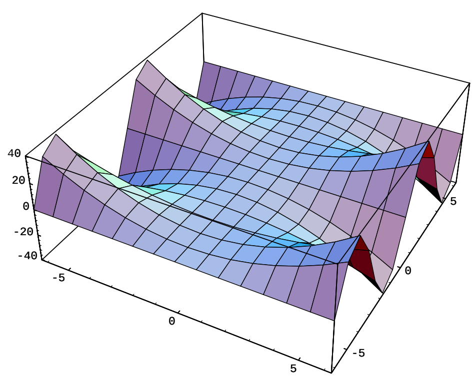
2. The *length* of the gradient is the *rate* of most rapid increase. For example, consider a function of two variables $z = f(x, y)$. We compared ∇f to the needle of a compass which always points in the "uphill direction". Property 2 says that the *length* of the needle changes from point to point, so that it's always equal to the magnitude of the "uphill slope".

You get the largest positive rate of change going uphill and the largest negative rate of change going downhill. If you walk along the hill "sideways", your altitude doesn't change and your rate of change is 0. Since "sideways" is perpendicular to uphill-dowhill, and since a line of constant altitude is a level curve, we have:

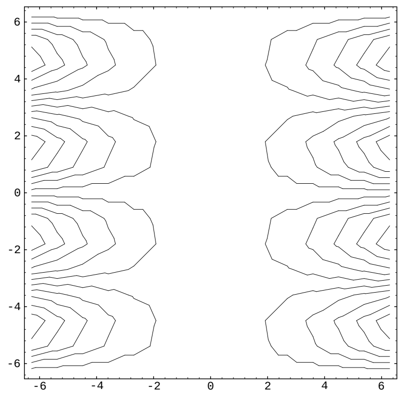
3. The gradient vector at a point is perpendicular to the level curve (or level surface, or level hypersurface, or ...) which passes through the point.

For a function of two variables $z = f(x, y)$, the gradient is a 2-dimensional vector which lies in the x-y-plane together with the level curves. There is a gradient vector at each point in the plane (which is why we say that ∇f is a **vector field**). A level curve represents a line of constant altitude. Since the gradient points "uphill", it is reasonable that it should be perpendicular to the level curve.

Example. Here is the graph of $z = x^2 \sin y$:



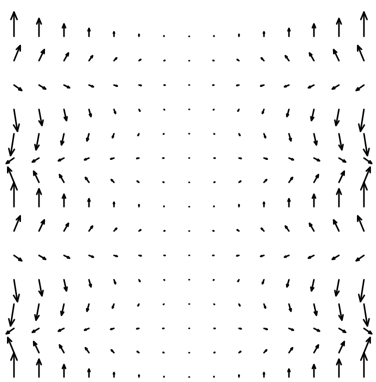
Here is a picture of the level curves:



The gradient field is

$$\nabla f = \langle 2x \sin y, x^2 \cos y \rangle.$$

Here is the gradient field:



Notice that the vectors in the gradient field are (approximately) perpendicular to the level curves. □

Example. Find the gradient of $f(x, y, z) = \dfrac{1}{\sqrt{x^2 + y^2 + z^2 + 1}}$.

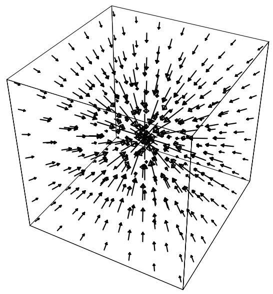
$$\nabla f = \left\langle \frac{-x}{(x^2 + y^2 + z^2 + 1)^{3/2}}, \frac{-y}{(x^2 + y^2 + z^2 + 1)^{3/2}}, \frac{-z}{(x^2 + y^2 + z^2 + 1)^{3/2}} \right\rangle.$$

We can write this in the form

$$\nabla f = \frac{-1}{(x^2 + y^2 + z^2 + 1)^{3/2}} \langle x, y, z \rangle.$$

$\langle x, y, z \rangle$ is the **radial vector** from the origin $(0, 0, 0)$ to the point (x, y, z). In this example, ∇f always points *inward* toward the origin. ("Inward" because of the -1.)

Here's a picture of the gradient vector field:



The vectors do indeed point inward toward the origin. □

As an application of Property 3, we can find the equation of the **tangent plane to a surface**.

Example. Find the tangent plane to $x^2 - y^2 + 2yz + z^5 = 6$ at the point $(2, 1, 1)$.

To write down the equation of a plane, we need a point on the plane and a vector perpendicular to the plane.

The point of tangency $(2, 1, 1)$ will serve as the point on the plane.

A vector perpendicular to the tangent plane is also perpendicular to the surface. Such a vector is called a **normal vector** for the surface. There are infinitely many normal vectors to a surface at a given point, since we can take any one normal vector and scale it up or down.

You may know that a normal vector to a surface $z = f(x, y)$ is given by

$$\vec{N} = \left\langle -\frac{\partial f}{\partial x}, -\frac{\partial f}{\partial y}, 1 \right\rangle.$$

We can't use that result here, because we can't solve for z in terms of x and y.

Instead, we use a little trick and a property of the gradient. Write $w = x^2 - y^2 + 2yz + z^5 - 6$. (Take the original surface and drag everything to one side of the equation.) The original surface is $w = 0$.

Now $w = 0$ is a level surface of w, and the gradient ∇w is perpendicular to level surfaces of w. Hence, ∇w must be perpendicular to the original surface.

We compute the gradient and plug in $(2, 1, 1)$:

$$\nabla w = \langle 2x, -2y + 2z, 2y + 5z^4 \rangle, \quad \nabla w(2, 1, 1) = \langle 4, 0, 7 \rangle.$$

The vector $\langle 4, 0, 7 \rangle$ is perpendicular to the tangent plane. Hence, the plane is

$$4(x - 2) + 0 \cdot (y - 1) + 7(z - 1) = 0, \quad \text{or} \quad 4x + 7z = 15. \ \square$$

Problems on gradients and directional derivatives.

1. Let $f(x, y) = \dfrac{x^2}{x+y}$.

(a) Find $\nabla f(1, -2)$.

(b) At what rate is f changing at the point $(1, -2)$ in the northeast direction?

(c) Find a unit vector which points in the steepest uphill direction at the point $(1, -2)$.

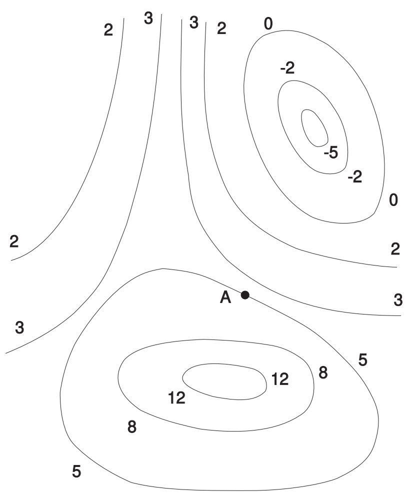
2. Consider the surface defined by the equation $x^2y + xz^2 = z^3 + 3x - 8$.

(a) Show that the point $(1, -1, 2)$ lies on the surface.

(b) Find the equation of the tangent plane to the surface at the point $(1, -1, 2)$.

(c) Find the equation of the line which passes through $(1, -1, 2)$ and is perpendicular to the surface at that point.

3. The picture below shows some level curves for a function $z = f(x, y)$.



(a) At what point does there appear to be a local maximum?

(b) Draw a vector at the point A which points in the direction of the gradient vector at A.

4. The rate of change of $f(x, y)$ at $(1, -1)$ is 2 in the direction *toward* $(5, -1)$ and is $\dfrac{6}{5}$ in the direction of the vector $\langle -3, -4 \rangle$. Find $\nabla f(1, -1)$.

Solutions to the problems on gradients and directional derivatives.

1. Let $f(x,y) = \dfrac{x^2}{x+y}$.

(a) Find $\nabla f(1,-2)$.

$$\nabla f = \langle \frac{(x+y)(2x)-(x^2)(1)}{(x+y)^2}, \frac{-x^2}{(x+y)^2} \rangle = \langle \frac{x^2+2xy}{(x+y)^2}, \frac{-x^2}{(x+y)^2} \rangle, \quad \nabla f(1,-2) = \langle -3,-1 \rangle. \square$$

(b) At what rate is f changing at the point $(1,-2)$ in the northeast direction?

The vector $\vec{v} = \langle 1,1 \rangle$ points to the northeast, as you can see by drawing a picture. Since $|\langle 1,1 \rangle| = \sqrt{2}$, we have

$$Df_{\vec{v}}(1,-2) = \nabla f(1,-2) \cdot \frac{\vec{v}}{|\vec{v}|} = \langle -3,-1 \rangle \cdot \frac{1}{\sqrt{2}} \langle 1,1 \rangle = -2\sqrt{2} \approx 2.82843. \square$$

(c) Find a unit vector which points in the steepest uphill direction at the point $(1,-2)$.

The gradient $\nabla f(1,-2) = \langle -3,-1 \rangle$ points in the direction of most rapid increase ("steepest uphill direction") at the point $(1,-2)$. Since $|\langle -3,-1 \rangle| = \sqrt{10}$, the unit vector in this direction is $\dfrac{1}{\sqrt{10}} \langle -3,-1 \rangle$. $\square$

2. Consider the surface defined by the equation $x^2y + xz^2 = z^3 + 3x - 8$.

(a) Show that the point $(1,-1,2)$ lies on the surface.

If we substitute $(1,-1,2)$ into the surface equation, we obtain $(1)(-1) + (1)(4) = 8 + (3)(1) - 8$, or $8 = 8$.

The point satisfies the surface equation, so the point is on the surface. $\square$

(b) Find the equation of the tangent plane to the surface at the point $(1,-1,2)$.

Rewrite the surface equation in the form $x^2y + xz^2 - z^3 - 3x + 8 = 0$. and set $w = x^2y + xz^2 - z^3 - 3x + 8$. Then our surface is the level surface $w = 0$. $\nabla w(1,-1,2)$ is perpendicular to the level surface $w = 0$, so it is perpendicular to our surface.

We have

$$\nabla w = \langle 2xy + z^2 - 3, x^2, 2xz - 3z^2 \rangle, \quad \text{and} \quad \nabla w(1,-1,2) = \langle -1,1,-8 \rangle.$$

The vector $\langle -1,1,-8 \rangle$ is perpendicular to the tangent plane, so the equation of the tangent plane is

$$(-1)(x-1) + (1)(y+1) + (-8)(z-2) = 0, \quad \text{or} \quad x - y + 8z = 18. \square$$

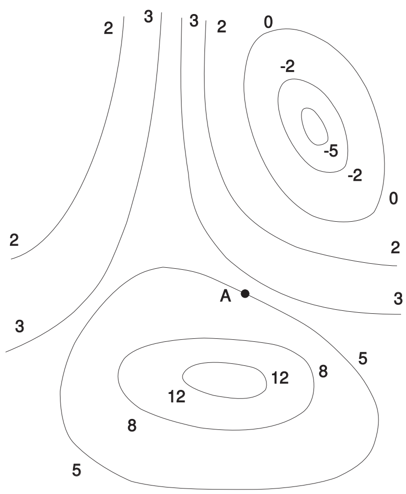
(c) Find the equation of the line which passes through $(1,-1,2)$ and is perpendicular to the surface at that point.

The vector $\nabla w(1,-1,2) = \langle -1,1,-8 \rangle$ is perpendicular to the surface at $(1,-1,2)$, so the line

$$x - 1 = -t, \quad y + 1 = t, \quad z - 2 = -8t$$

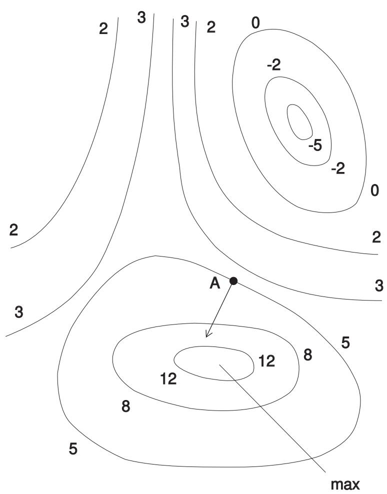
passes through $(1,-1,2)$ and is perpendicular to the surface at that point. $\square$

3. The picture below shows some level curves for a function $z = f(x,y)$.



(a) At what point does there appear to be a local maximum?

The position of the local max is indicated on the picture below:



(b) Draw a vector at the point A which points in the direction of the gradient vector at A.

The gradient vector points in the direction of most rapid increase, and is perpendicular to the level curve passing through a point. Hence, the vector you draw should be perpendicular to the level curve which passes through A, and should point in the general direction of the lines of higher altitude. An appropriate vector is shown on the picture above. □

4. The rate of change of $f(x, y)$ at $(1, -1)$ is 2 in the direction *toward* $(5, -1)$ and is $\frac{6}{5}$ in the direction of the vector $\langle -3, -4\rangle$. Find $\nabla f(1, -1)$.

The direction from $(1, -1)$ *toward* the point $(5, -1)$ is given by the vector $\langle 4, 0\rangle$. This vector has length 4, so

$$2 = \nabla f(1, -1) \cdot \frac{\langle 4, 0\rangle}{4} = \langle f_x, f_y\rangle \cdot \frac{\langle 4, 0\rangle}{4} = f_x.$$

The vector $\langle -3, -4\rangle$ has length 5, so

$$\frac{6}{5} = \nabla f(1, -1) \cdot \frac{\langle -3, -4\rangle}{5} = \langle f_x, f_y\rangle \cdot \frac{\langle -3, -4\rangle}{5} = -\frac{3}{5} f_x - \frac{4}{5} f_y.$$

Thus, $6 = -3f_x - 4f_y$.

We have two equations involving f_x and f_y. If we solve simultaneously, we obtain $f_x = 2$ and $f_y = -3$. Hence, $\nabla f(1, -1) = \langle 2, -3 \rangle$. $\square$